

# Draft Genome Sequence of a *Legionella pneumophila* Serogroup 4 Strain Causing Legionellosis

Akira Okamoto,<sup>a,b</sup> Hiromichi Lee,<sup>b</sup> Mitsutaka Yabutani,<sup>c</sup> Keiko Yamada,<sup>b</sup> Michio Ohta<sup>b,d</sup>

Department of School Health Sciences, Aichi University of Education, Kariya, Japan<sup>a</sup>; Department of Bacteriology, Nagoya University Graduate School of Medicine, Showa-ku, Nagoya, Japan<sup>b</sup>; Nagoya City Public Health Research Institute, Mizuho-ku, Nagoya, Japan<sup>c</sup>; School of Nursing, Sugiyama Jogakuen University, Chikusa-ku, Nagoya, Japan<sup>d</sup>

Here, we report the draft genome sequence of the *Legionella pneumophila* Nagoya-1 strain, serogroup 4, which was isolated from a clinical sample from a patient with legionellosis. Several virulence-associated genes, including those encoding the type IV (Dot/Icm) secretion system and effector proteins, were highly conserved.

Received 27 May 2014 Accepted 6 June 2014 Published 19 June 2014

Citation Okamoto A, Lee H, Yabutani M, Yamada K, Ohta M. 2014. Draft genome sequence of a *Legionella pneumophila* serogroup 4 strain causing legionellosis. *Genome Announc.* 2(3):e00602-14. doi:10.1128/genomeA.00602-14.

Copyright © 2014 Okamoto et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Akira Okamoto, okamoto@uecc.aichi-edu.ac.jp.

*L*egionella pneumophila is the causative agent of legionellosis, including pneumonia and Pontiac fever, and it is transferred by waterborne aerosolized bacteria rather than human-to-human interaction. *L. pneumophila* strains are categorized by serogroups, on the basis of the lipopolysaccharide cell surface structure. Serogroup 1 has been well investigated, because it is the predominant serogroup among clinical isolates (1, 2). Although other serogroups are often isolated, few have been investigated (3, 4).

Here, we report a draft genome sequence of *L. pneumophila* serogroup 4 strain Nagoya-1. This strain was isolated in 2000 from a patient with legionellosis in a hospital in Japan. Simultaneously, another isolate of *L. pneumophila* serogroup 4 was obtained from a bathroom in the hospital. The two isolates yielded comparable pulsed-field gel electrophoresis patterns.

Genomic DNA was extracted from an overnight culture of *L. pneumophila* Nagoya-1 by using a Wizard genomic DNA purification kit (Promega, Madison, WI). A DNA library was prepared for sequencing with a DNA sample prep kit (Illumina, San Diego, CA), according to the supplied protocol. Single-end sequencing with a read length of 36 bp was performed using a genome analyzer system (Illumina). We obtained 6,130,067 reads that passed the quality test, and the total amount of read data comprises 220,682,412 bp, representing approximately 60-fold coverage of the genome. The read sequences were compiled using Edena version 2.1.1 software (5). The assembled genome consists of 426 contigs, with a total length of 3,363,464 bp and 38.2% average G+C content. The genome sequence was annotated by the Rapid Annotations using Subsystems Technology (RAST) server (6). In all, 3,221 features were identified in 426 contigs, consisting of 3,185 predicted coding sequences, 33 tRNA genes, and 3 rRNAs. tRNA predictions were confirmed by the tRNAscan-SE and RNAmmer servers (7, 8), revealing results consistent with those from the RAST server. The contigs were mapped as a query to the reference genome, *L. pneumophila* strain Philadelphia 1 (serogroup 1, 3,397,754 bp), using the MUMmer (version 3.0) pro-

gram; however, only 701,725 bp (20.7% of the reference genome) was covered with the contigs.

Genes encoding the type IV (Dot/Icm) secretion system and several effector proteins were annotated. A homology search analysis was performed with the nucleotide sequences of genes encoding the type IV secretion system and effector proteins from the Philadelphia 1 strain as queries, and contigs of Nagoya-1 were used as the database. All of the genes in the Philadelphia 1 dot/icm loci were found in the Nagoya-1 genome, with nucleotide sequence identities of 84 to 100%. Genes encoding several effectors, including *legL3*, *legLC8*, *lepA*, *lepB*, *lepB-1*, *shdA*, *sidF*, *vipA*, and *vipF*, were confirmed, with identities of 95.0 to 99.5%.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited in the DNA Data Bank of Japan (DDBJ), European Molecular Biology Laboratory, and GenBank under the accession numbers [BAZA01000001](#) to [BAZA01000426](#). The run data generated from sequencing were deposited in the DRA database in the DDBJ under the accession number [DRA002158](#).

## ACKNOWLEDGMENTS

This study was supported by a grant from the Hori Sciences and Arts Foundation and by a Grant-in-Aid for Research from Nagoya University.

We thank Hiromasa Ono and Hidemasa Bono on TogoTV (9) of the Database Center for Life Science, the Research Organization of Information and Systems (DBCLS), and Norichika Ogata for the suggestion to perform *in silico* analysis.

## REFERENCES

- Chien M, Morozova I, Shi S, Sheng H, Chen J, Gomez SM, Asamani G, Hill K, Nuara J, Feder M, Rineer J, Greenberg JJ, Steshenko V, Park SH, Zhao B, Teplitskaya E, Edwards JR, Pampou S, Georgiou A, Chou IC, Iannuccilli W, Ulz ME, Kim DH, Geringer-Sameth A, Goldsberry C, Morozov P, Fischer SG, Segal G, Qu X, Rzhetsky A, Zhang P, Cayanan E, De Jong PJ, Ju J, Kalachikov S, Shuman HA, Russo JJ. 2004. The genomic sequence of the accidental pathogen *Legionella pneumophila*. *Science* 305:1966–1968. <http://dx.doi.org/10.1126/science.1099776>.

2. Cazalet C, Rusniok C, Brüggemann H, Zidane N, Magnier A, Ma L, Tichit M, Jarraud S, Bouchier C, Vandenesch F, Kunst F, Etienne J, Glaser P, Buchrieser C. 2004. Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat. Genet.* 36:1165–1173. <http://dx.doi.org/10.1038/ng1447>.
3. Khan MA, Knox N, Prashar A, Alexander D, Abdel-Nour M, Duncan C, Tang P, Amatullah H, Dos Santos CC, Tijet N, Low DE, Pourcel C, Van Domselaar G, Terebiznik M, Ensminger AW, Guyard C. 2013. Comparative genomics reveal that host-innate immune responses influence the clinical prevalence of *Legionella pneumophila* serogroups. *PLoS One* 8:e67298. <http://dx.doi.org/10.1371/journal.pone.0067298>.
4. Amaro F, Gilbert JA, Owens S, Trimble W, Shuman HA. 2012. Whole-genome sequence of the human pathogen *Legionella pneumophila* serogroup 12 strain 570-CO-H. *J. Bacteriol.* 194:1613–1614. <http://dx.doi.org/10.1128/JB.06626-11>.
5. Hernandez D, François P, Farinelli L, Osterås M, Schrenzel J. 2008. *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res.* 18:802–809. <http://dx.doi.org/10.1101/gr.072033.107>.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
7. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964. <http://dx.doi.org/10.1093/nar/25.5.0955>.
8. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
9. Kawano S, Ono H, Takagi T, Bono H. 2012. Tutorial videos of bioinformatics resources: online distribution trial in Japan named TogoTV. *Brief. Bioinform.* 13:258–268. <http://dx.doi.org/10.1093/bib/bbr039>.